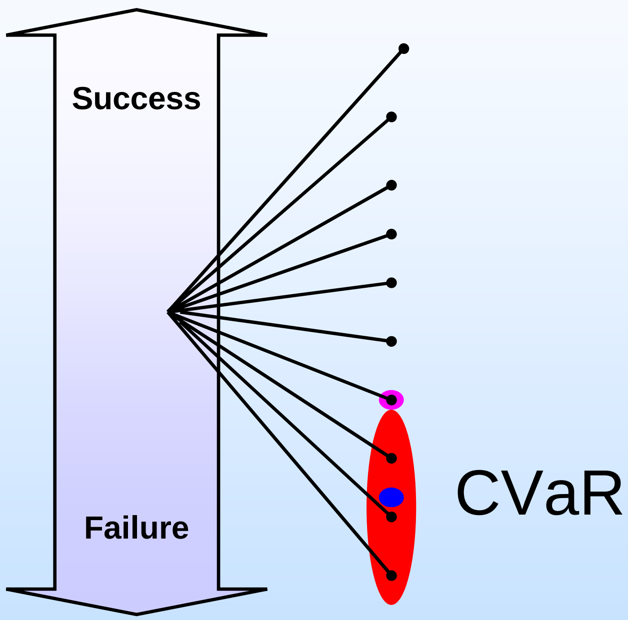
Combining Model Analysis and Experimental Test Data for Optimal Determination of Failure Allowables



Stan Uryasev

University of Florida

uryasev@ufl.edu
www.ise.ufl.edu/uryasev

and

American Optimal Decisions

Joint presentation with A.A. Trindade

This effort was jointly accomplished by the Boeing led team and the United States Government under the guidance of NAVAIR. This work was funded by DARPA/DSO and administered by NAVAIR under Technology Investment Agreement N00421-01-3-0098. The program acknowledges the guidance and support of Dr. Leo Christodoulou of DARPA/DSO for this effort. The technical monitor for the program is Dr. Ray Meilunas of NAVAIR.

Report Documentation Page	Form Approved OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
2004	**N/A**	**-**

4. TITLE AND SUBTITLE **Combining Model Analysis and Experimental Test Data for Optimal Determination of Failure Allowables**	5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **University of Florida**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR’S ACRONYM(S) 11. SPONSOR/MONITOR’S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**UU**	**38**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

Objectives

- Pooling various sources of information of possibly different pedigree: models, experiments and expert opinions
- Methodology for estimating percentiles of failure distributions and allowables
- Minimizing amount of data needed for certification process
- Accounting for various sources of uncertainty
- Demonstration of the approach with case studies
- Validation in controlled statistical environment

Accomplished Tasks

- **Developed factor model for <u>direct</u> estimation of percentiles using:**

  - various sources of information (models, experiments, expert opinions); it is possible to quantify value of different inputs

  - statistical characteristics: mean, st.dev., deviation CVaR, etc.

- **simple, clear, computationally effective methodology enables pooling data across**

  - many individual materials: relatively small requirements to size of the datasets

  - various experiment setups: crediting simple experiments to more sophisticated (expensive) ones

Accomplished Tasks (cont'd)

- **Developed CVaR statistical techniques for optimal estimating weighting coefficients of factor model and confidence intervals (A-basis and B-basis)**

  - percentile/CVaR regression approach was specially designed for estimating percentiles and confidence intervals (A and B-basis)

  - CVaR deviation measure has exceptional mathematical and computational qualities: coherent measure (convexity, etc.)

  - NO distribution assumptions (such as normality, ...)

  - approach is based on linear programming: very large datasets, stability of results, high speed of calculations

- **CVaR statistical technique for combining various modeling and experimental inputs is new; it was developed in the framework of AIM-C project**

Accomplished Tasks (Cont'd)

- **Two case studies for open-hole coupon dataset: estimation of 10th percentile and B-basis (failure data + 3 models)**

  - In-sample calculations with CVaR regression provided correct estimates of percentiles of failure load distributions.

  - Models via CVaR regression provide plausible percentile estimates and B-basis, even in the absence of any experimental test information.

  - Out-of-sample 10th percentile estimates based on Model 1 individually, and on Model 1 plus 5 test points, are close to their true values. B-basis values are also close to nominal values based on actual experiments.

  - Benefits of combining models for predicting percentiles were quantified

  - Benefits of combining models and experimental data were evaluated

Accomplished Tasks (Cont'd)

- **Case study with Monte-Carlo simulated data**

  - minimal number of stacking sequences needed to populate the CVaR regression model and sensitivity to this number

  - minimal number of experimental datapoints per stacking sequence and sensitivity to this number

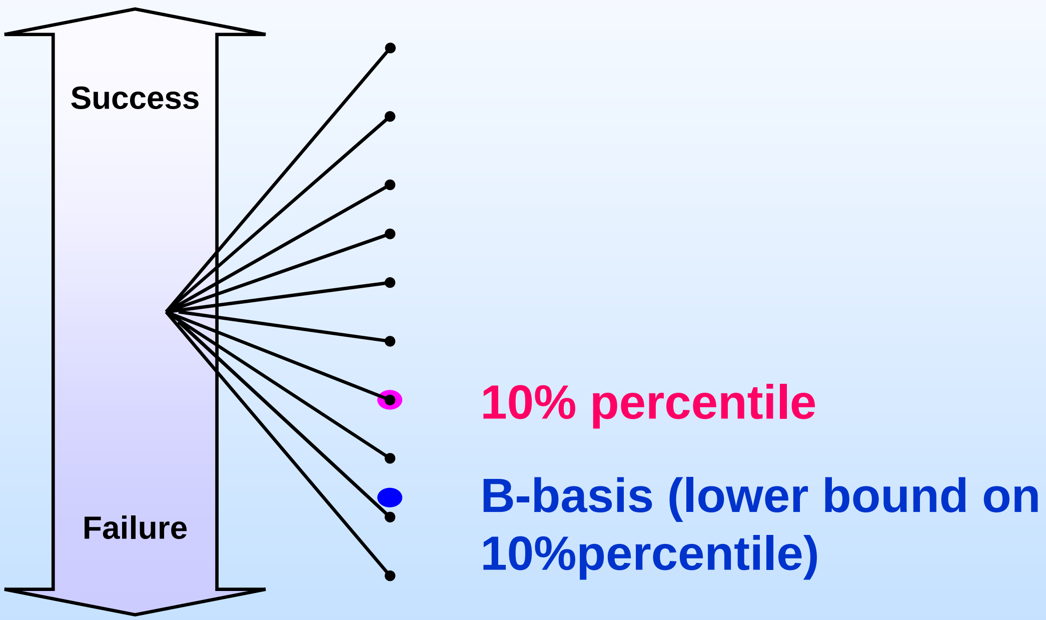
  - CVaR error (which is observer) versus "true" error in percentiles which can not be observed

  - sensitivity of the approach to errors in models

- **Our investigations provide compelling evidence that the methodology we are developing can integrate modeling and experimental data and radically reduce overall testing cost**

Percentiles and Allowables

Failure load distribution



Factor Models: Percentile and CVaR Regression

factors $X_1, ..., X_q$ **from various sources of information**

failure load Y

$Y = c_0 + c_1 X_1 +, ..., + c_q X_q + \varepsilon$, **where** ε **is an *error* term**

$c_0 + c_1 X_1 +, ..., + c_q X_q$ **= direct estimator of percentile with α confidence**

Percentile regression (Koenker and Basset (1978))

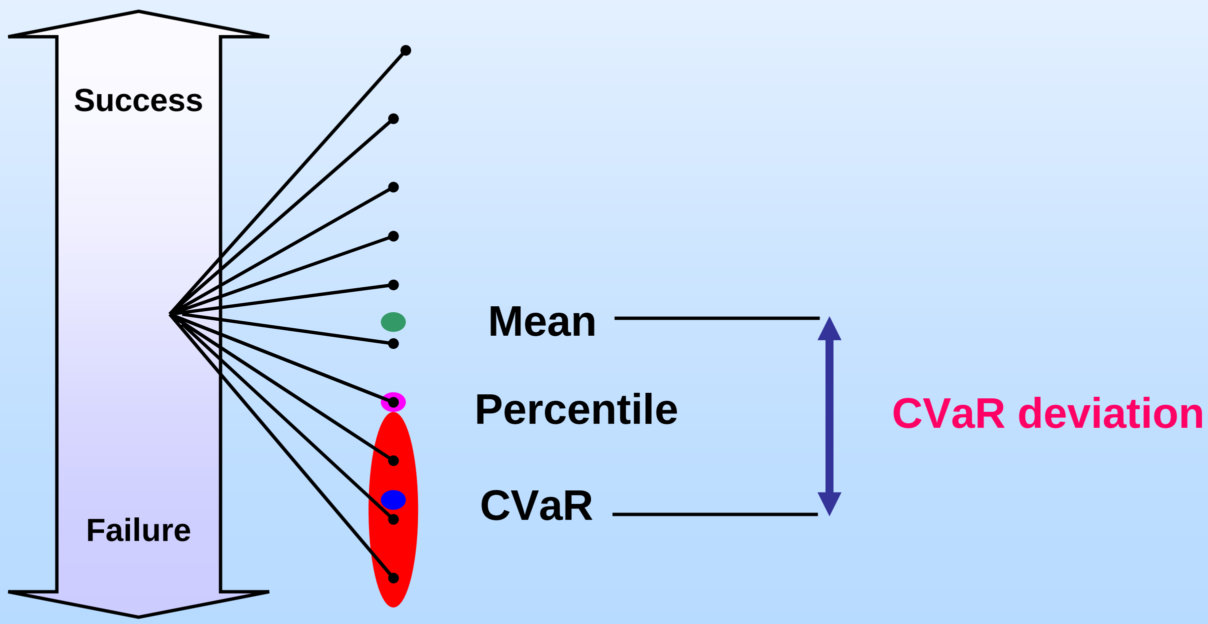
CVaR regression (Rockafellar, Uryasev, Zabarankin (2003))

Percentile Error Function and CVaR Deviation

Statistical methods based on asymmetric percentile error functions: $E[(1-\alpha)(-\varepsilon^{-})+\alpha\varepsilon^{+}]$

ε^{+} **= positive part of error**

ε^{-} **= negative part of error**



Case Study 1: Dataset

- **Two rows of data (totally 35 rows, test type = 13, test type 2 = 14, test type 3 = 4, test type 4 = 4)**

	W/D	**Test Data**	**Prediction Model 1 upper estimate**	**Prediction midpoint**	**Prediction Model 1 lower estimate**
	6	137.34	115.6		109
		118.96			
		119.41			
		110.64			
		103.67			
	mean	*118*	mean	*112.3*	
	6	48.3	56.64		51.68
		48.36			
		49.54			
		49.88			
		46.02			
	mean	*48.42*	mean	*54.16*	

How It Works?

μ = (115.6+109)/2

σ = (115.6-109)/2

137.34	115.6	109
118.96		
119.41		
110.64		
103.67		
48.3	56.64	51.68
48.36		
49.54		
49.88		
46.02		

μ	σ	137.34
μ	σ	118.96
μ	σ	119.41
μ	σ	110.64
μ	σ	103.67
μ	σ	48.3
μ	σ	48.36
μ	σ	49.54
μ	σ	49.88
μ	σ	46.02

$(c_0 + c_1\,\mu + c_2\sigma) - 137.34 =$ *error* 1

$(c_0 + c_1\,\mu + c_2\sigma) - 118.96 =$ *error* 2

...

$(c_0 + c_1\,\mu + c_2\sigma) - 48.3 =$ *error* 1

$(c_0 + c_1\,\mu + c_2\sigma) - 48.36 =$ *error* 2

...

$$\min CVaR^{\Delta}_{0.9}[-error]$$ => coefficients (c_0, c_1, c_2) for 10%

$$\min CVaR^{\Delta}_{0.8}[-error]$$ => coefficients (c_0, c_1, c_2) for 20%

- **Linear Programming, large dimensions**

CVaR Regression: Model 1 predictions

- **Factors (midpoint and width) generated by Model 1**

$$\mu^1 = (115.6 + 109)/2 \quad , \qquad \sigma^1 = (115.6 - 109)/2$$

$$\mu^2 = (56.64 + 51.68)/2 \; , \qquad \sigma^2 = (56.64 - 51.68)/2$$

$$\cdots$$

- **90% estimator obtained by CVaR regression**

$$y^i_{90\%} = \text{-}6.870 + 1.138\ \mu^i + 0.788\ \sigma^i$$

$$CVaR^{\Delta}_{0.9}\ [\varepsilon] = 13.337$$

- **10% estimator obtained by CVaR regression**

$$y^i_{10\%} = \text{-}2.427 + 0.974\ \mu^i\ \text{-}\ 1.008\ \sigma^i$$

$$CVaR^{\Delta}_{0.9}\ [-\varepsilon] = 9\, CVaR^{\Delta}_{0.1}[\varepsilon] \ = 8.905$$

CVaR Regression, In-Sample Experiments Model 1 + 5 measurements

- **Factors calculated with experimental data**

m^i **= estimates of mean for material *i* based on experimental data**
s^i **= estimate of standard deviation for material *i* based on experimental data**

- **10% estimator obtained by CVaR regression**

$$y^i_{10\%} = 0.430 - 0.015\,\mu^i - 0.075\,\sigma^i + 1.013 m^i - 1.129\, s^i$$

$$CVaR^{\Delta}_{0.9}[-\varepsilon] = 9\,CVaR^{\Delta}_{0.1} = 4.901$$

- **CVaR error (4.901) based on 5 experiments and Model 1 is significantly lower compared to CVaR error (8.905) based only on Model 1**

Out-of-sample Calculations: Raw # 11

- **Stacking sequence, raw 11, contains 294 measurements. It is considered for out-of sample verification of the approach.**

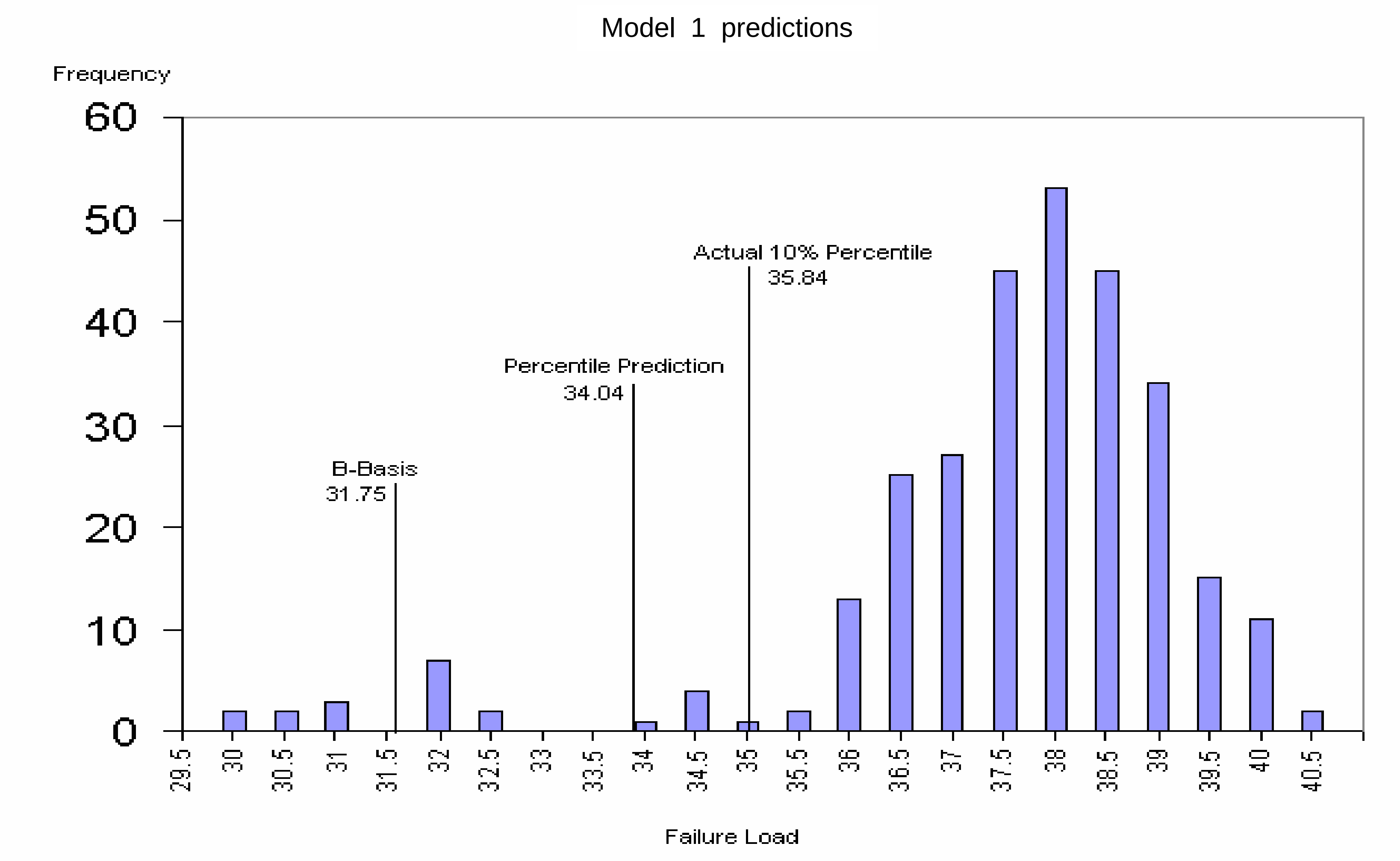
- **Model 1 lower and upper bounds equal 37.86 and 39.75, accordingly.**

- **CVaR regression based on Model 1 prediction is used to estimate 10% and B-basis for raw 11. B-basis is lower than 10% actual percentile calculated with 294 measurements (see the following graph).**

- **5,000 random samples of 5 measurements were generated. In 78% cases B-basis based on these measurements was lower than 10% actual percentile.**

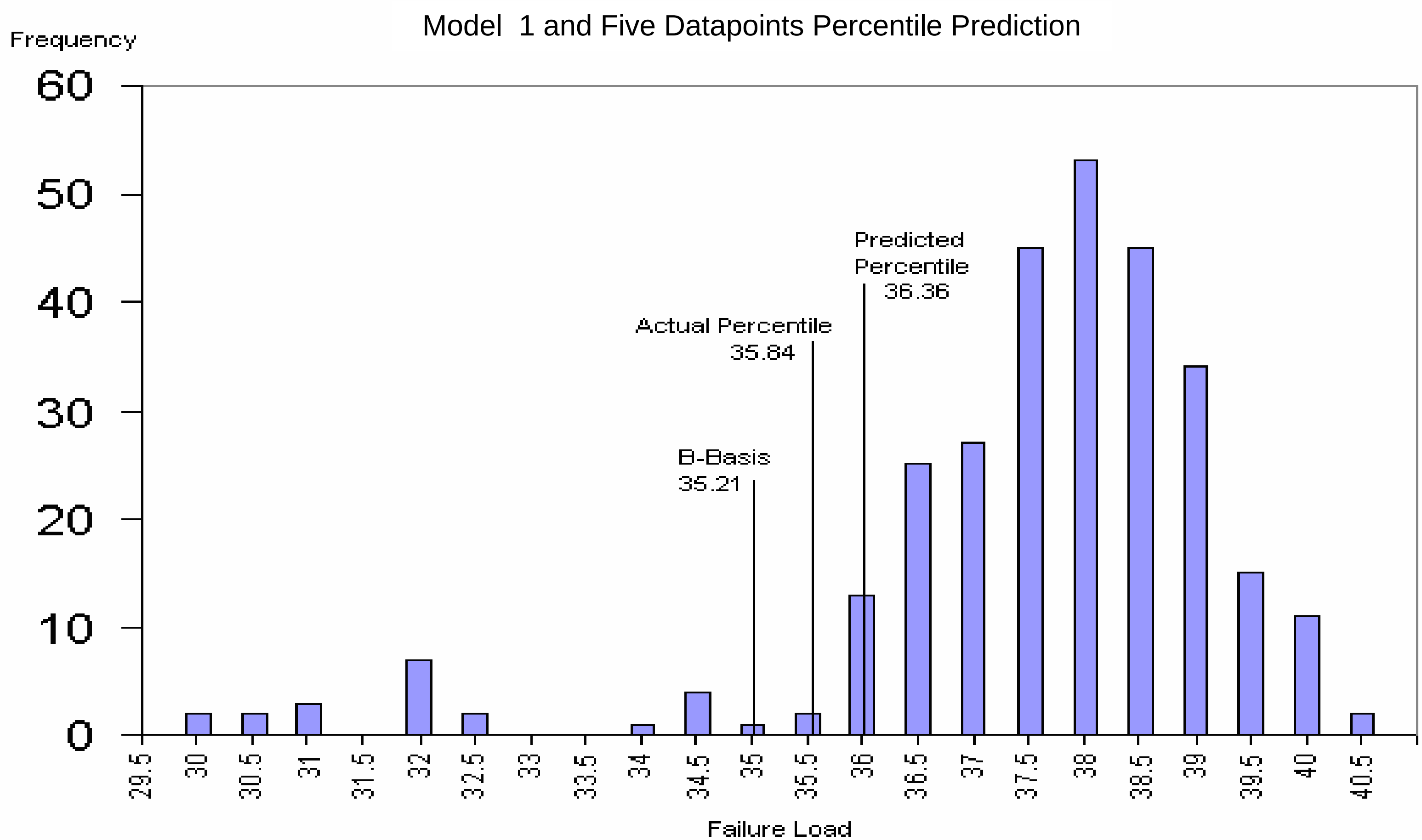
- **CVaR regression based on Model 1 prediction + 5 sample measurements {38.12, 37.214, 37.637, 37.707, 35.63} are used to estimate 10% and B-basis (see the graph).**

CVaR Regression: Model 1 predictions
10% and B-basis



Approved for Public Release, Distribution Unlimited

CVaR Regression: Model 1 +5 Measurements 10% and B-basis



Case Study 2: Dataset

- **One row of data (totally 28 rows, Test type 1 = 10, Test Type 2 = 10, Test Type 3 = 4, Test Type 4 = 4)**
- **Keep only 19 stacking sequences having at least 5 measurements**
- **Symmetric dataset was created with only 5 measurements per sequence**

			Model 1 (2 levels)		Model 2 (100 samples)			Model 3 (100 outcomes)		
	Type	Data	UL	LL		Average	Sigma		Average	Sigma
1	**Test Type 1**	**38.65**	**43.83**	**37.95**		**42.90**	**1.00**		**37.40**	**1.03**
		39.02			**10%**	**41.46**		**10%**	**35.95**	
		39.78			**90%**	**44.17**		**90%**	**38.76**	
		38.46								
		40.7								

Case Study 2: Objectives

- **Evaluation of**
  - **predictive capabilities of models for estimating percentiles and allowables**
  - **benefits of combining of models**
  - **benefits of combining models and actual experimental data**
  - **savings in number of actual experiments (replaced by model predictions)**

- **Suggestions for models calibration and removing of outliers**

Predicting 10% : Model Ranking

- **M# = model # ; T# = # of measurements**

Setup	regression coefficients								CVaR
	mean	st.dev	mean	st.dev	mean	st.dev	mean	st.dev	
M1			1.098	- 4.303					16.86
M2					0.571	0.005			24.656
M3							0.594	- 0.314	27.161
M123			0.510	- 6.005	0.660	-1.243	0.0409	0.040	13.822
T5	1.000	-1.435							10.287

- ***Model 1 (UL and LL) :*** **lowest CVaR = 16.86 among three models**
- ***Model 2 (Monte Carlo):*** **CVaR = 24.656 is higher than in Model1, information on st.dev. is not used**
- ***Model 3 (Monte Carlo):*** **CVaR = 27.161 is higher than in Models 1, 2**
- ***Model 123 (combining 3 models):***
  - models combining reduces error: CVaR = 13.822
  - information on model 3 is not used, combining of Model 1 and 2 will give the same result as combining 3 models
- **Models can be used to predict percentile without any actual measurements.**
  CVaR (M123) = 13.822 is not far away from CVaR (T5) = 10.287 with 5 measurements per stacking sequence

Combining 3 Models and 1-5 Measurements

Setup	regression coefficients								
	mean	st.dev	mean	st.dev	mean	st.dev	mean	st.dev	CVaR
M123,T1	0.303		0.105	- 5.131	0.825	-1.058	- 0.081	0.072	12.609
M123,T2	0.437	0.215	-0.264	- 5.714	1.161	-1.029	- 0.179	0.091	12.365
M123,T3	0.624	-0.268	-0.136	- 3.881	0.713	-0.718	- 0.088	0.046	11.821
M123,T4	0.875	-0.876	-0.101	-1.640	0.333	-0.371	- 0.059	0.032	10.786
M123,T5	0.966	-1.428	0.155	0.163	- 0.110	-0.178	- 0.002	0.039	9.725
T5	1.000	-1.435							10.287

- **Models bring significant information (if only 1-3 measurements are available)**
- **Models give input similar to 2-3 measurements (if only 1-3 measurements are available)**
- **Model 3 can be excluded from consideration without significant losses of information**

Individual Models versus 5 Measurements

Setup	regression coefficients								
	mean	st.dev	mean	st.dev	mean	st.dev	mean	st.dev	CVaR
M1,T5	0.940	-1.457	0.069	- 0.257					10.205
M2,T5	1.012	-1.435			- 0.009	- 0.223			10.01
M3,T5	1.036	-1.436					- 0.043	0.032	10.136
T5	1.000	-1.435							10.287

- **Models do not make significant improvements in CVaR deviation if 5 measurements per stacking sequence is available**

Model Ranking: Removing Outliers

- **Combining models and data allows individual model calibration for each stacking sequence (not done at present time)**
- **Outlier = no measurements in interval {mean-sigma, mean+sigma}**
- **Removing of outliers is one of possible ways to calibrate models (screening out unacceptable models for specific stacking sequence)**
- **Model 1 (7 outliers), Model 2 (9 outliers), Model 3 (12 outliers)**

Setup	regression coefficients								
	mean	st.dev.	mean	st.dev.	mean	st.dev.	mean	st.dev.	CVaR
M1,T5	0.940	-1.457	0.069	- 0.257					10.205
$M1_O$ T5	0.983	- 1.546	0.021	0.022					10.044
M2,T5	1.012	-1.435			- 0.009	- 0.223			10.010
$M2_O$,T5	1.164	- 1.253			- 0.180	- 0.256			9.941
M3,T5	1.036	-1.436					- 0.043	0.032	10.136
T5	1.000	-1.435							10.287

Case Study 2: Summary

- **Models can be used to predict percentile without any actual measurements**
- **Combining of models improves predictive capabilities of the approach (Model 1 + Model 2 without Model 3)**
- **Models bring significant information if only 1-3 measurements for a new stacking sequence are available**
- **Additional model calibration is allowed when models are combined with data**

Estimation of 10% of Failure Distribution: Testing in Controlled Statistical Environment:

- **Minimal data requirements for CVaR regression approach?**
  - **how many stacking sequences is needed to calibrate the model?**
  - **how many actual measurements per stacking sequence?**

- **Sensitivity of the CVaR estimation procedure to data**
  - **true error versus CVaR error**
  - **dependence of error on amount of data**

- **Sensitivity to the quality of the model**
  - **perfect model**
  - **constant biases in model parameters**
  - **random errors in model parameters**

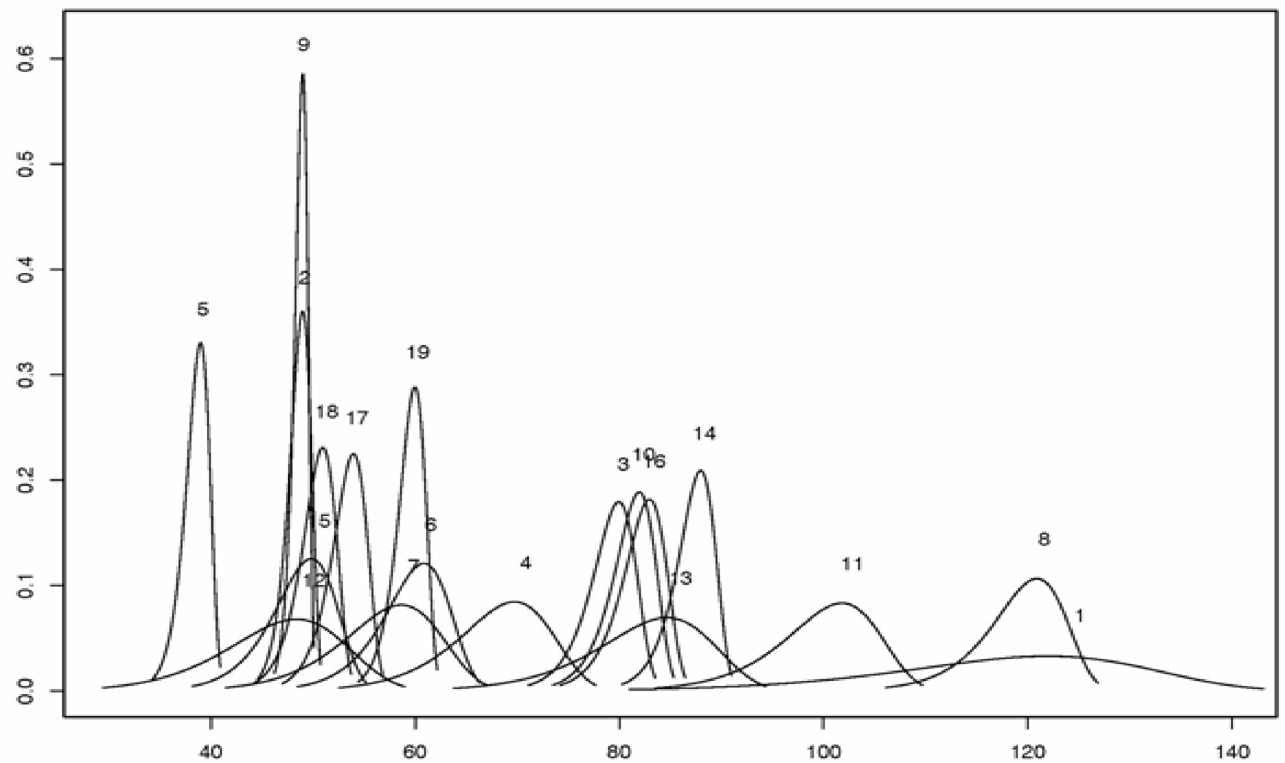
Lessons Learned from Monte Carlo Simulations

- **CVaR regression methodology is quite robust:**
  - **low sensitivity to systematic biases in scale and shape parameters of the model (completely insensitive to systematic bias in mean and st.dev.)**
- **Stable results for cases with more than 10 stacking sequences and more than 3 actual measurements per stacking sequence**
- **Relatively low sensitivity to number of stacking sequences (NO significant benefits to have more than 15 sequences)**
- **CVaR Deviation (which we can observe) is higher than the true error in percentile (which we can not observe), especially for cases with large number of actual measurements per stacking sequence**

Simulating Experimental and Modeling Data

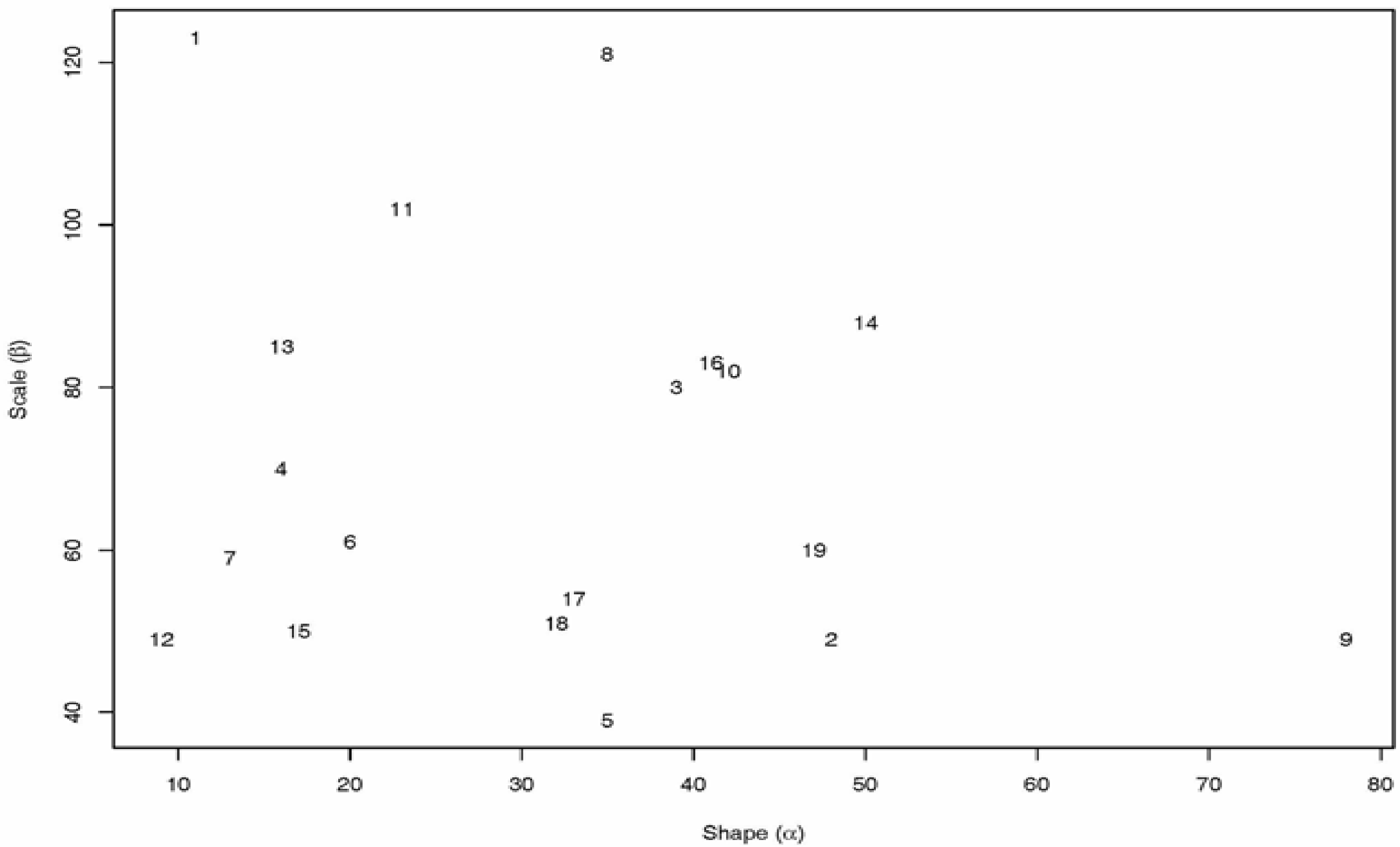
- Modeling experimental and test data by Weibull distributions: rich family of distributions with various shapes and mean values
- ***Two-Parameter Weibull Distribution:*** probability that an observation lies between *a* and *b* ($0 < a < b < \infty$) equals $e^{-(a/\alpha)^{\beta}} - e^{-(b/\alpha)^{\beta}}$, where α is ***scale parameter*** and β is ***shape parameter***
- Weibull distributions have been fit for 19 stacking sequences (having at least 5 actual measurements) with maximum likelihood approach
- Realistic range of parameters of Weibull distribution

Weibull Plots for 19 Stacking Sequences



Approved for Public Release, Distribution Unlimited

Range of Parameters: Weibull Distribution



Approved for Public Release, Distribution Unlimited

Assumptions on Models

- **Perfect Model**

  Model distribution coincides with true distribution

  (however, CVaR approach is nonparametric and it does not utilize information on distribution)

- **Constant Bias in Scale Parameter (similar to bias in mean value)**

  $\alpha \Rightarrow \alpha + 20$

- **Constant Bias in Shape Parameter (similar to bias in standard deviation)**

  $\beta \Rightarrow \beta + 20$

- **Random Bias in Scale Parameter**

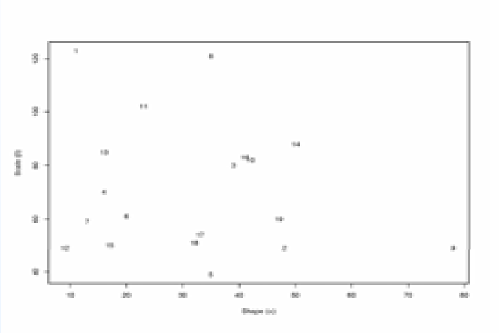
  $\alpha \Rightarrow \alpha + 20*N(0,1)$

- **Random Bias in Shape Parameter**

  $\beta \Rightarrow \beta + 20*N(0,1)$

Sampling Procedure

Step 1. **Sample uniformly *m* stacking sequences (points in the box of Weibull distributions)**



Step 2. **For each stacking sequence sample *n* “experimental” data points from “true” Weibull**

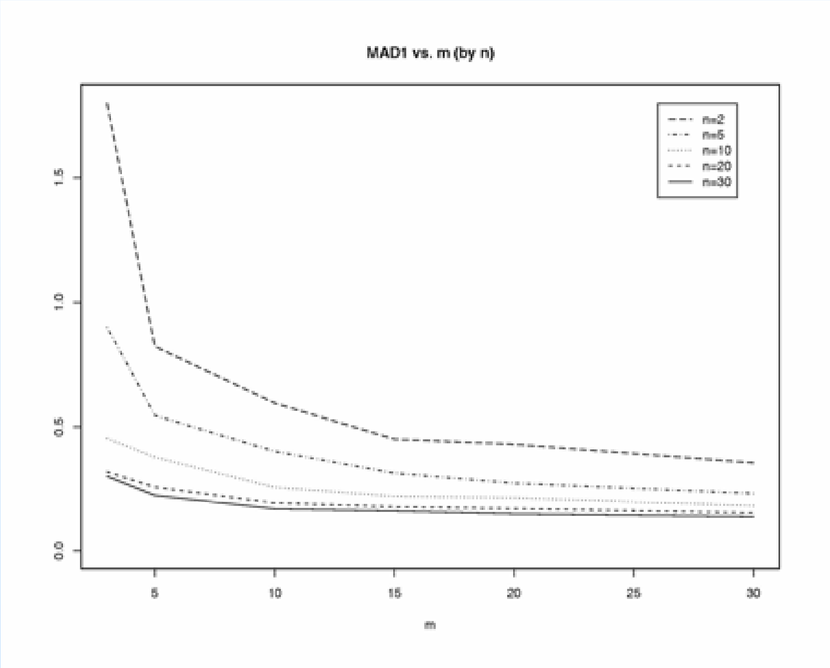
Step 3. **For each stacking sequence sample 100 “model” outcomes from “model” Weibull**

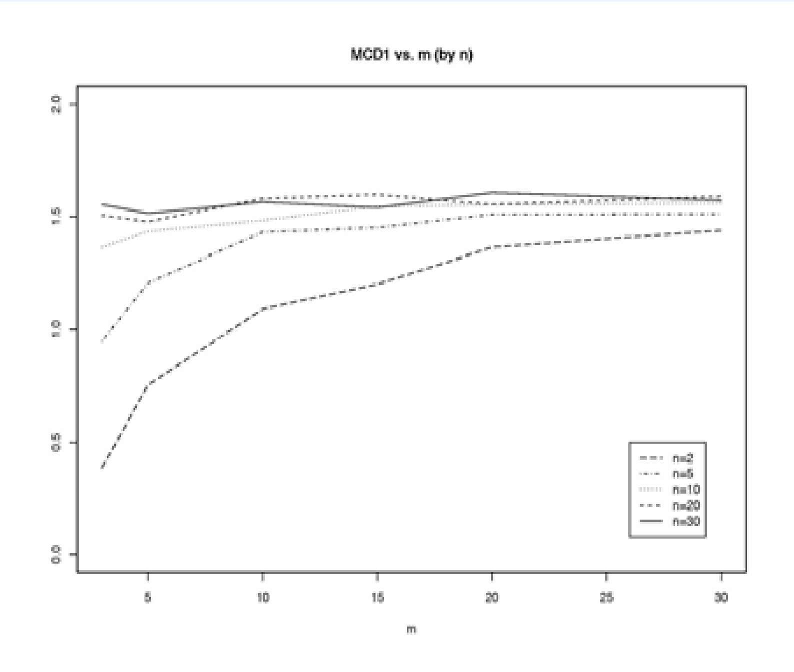
Step 4. **Use CVaR regression to predict percentile**

Repeat Steps 1- 4 100 times and calculate Mean Absolute Deviation (MAD) error and Mean CVaR Deviation over 100 runs

Perfect Model is Used for Predicting Percentiles

- **Low sensitivity to # of stacking sequences, NO significant benefits for m>15**
- **True error, Mean Absolute Deviation (MAD), ranges from 0.2 to 2**

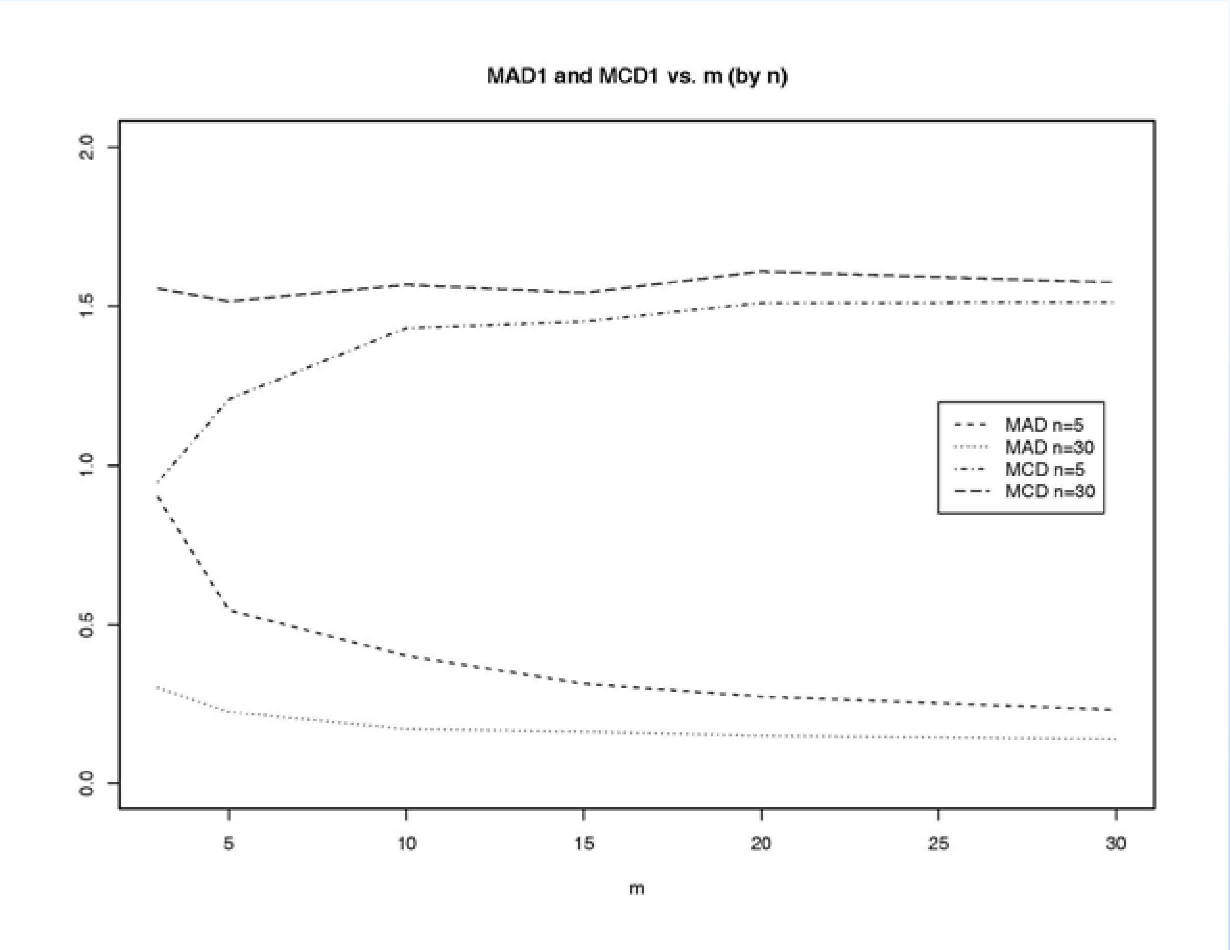




Approved for Public Release, Distribution Unlimited

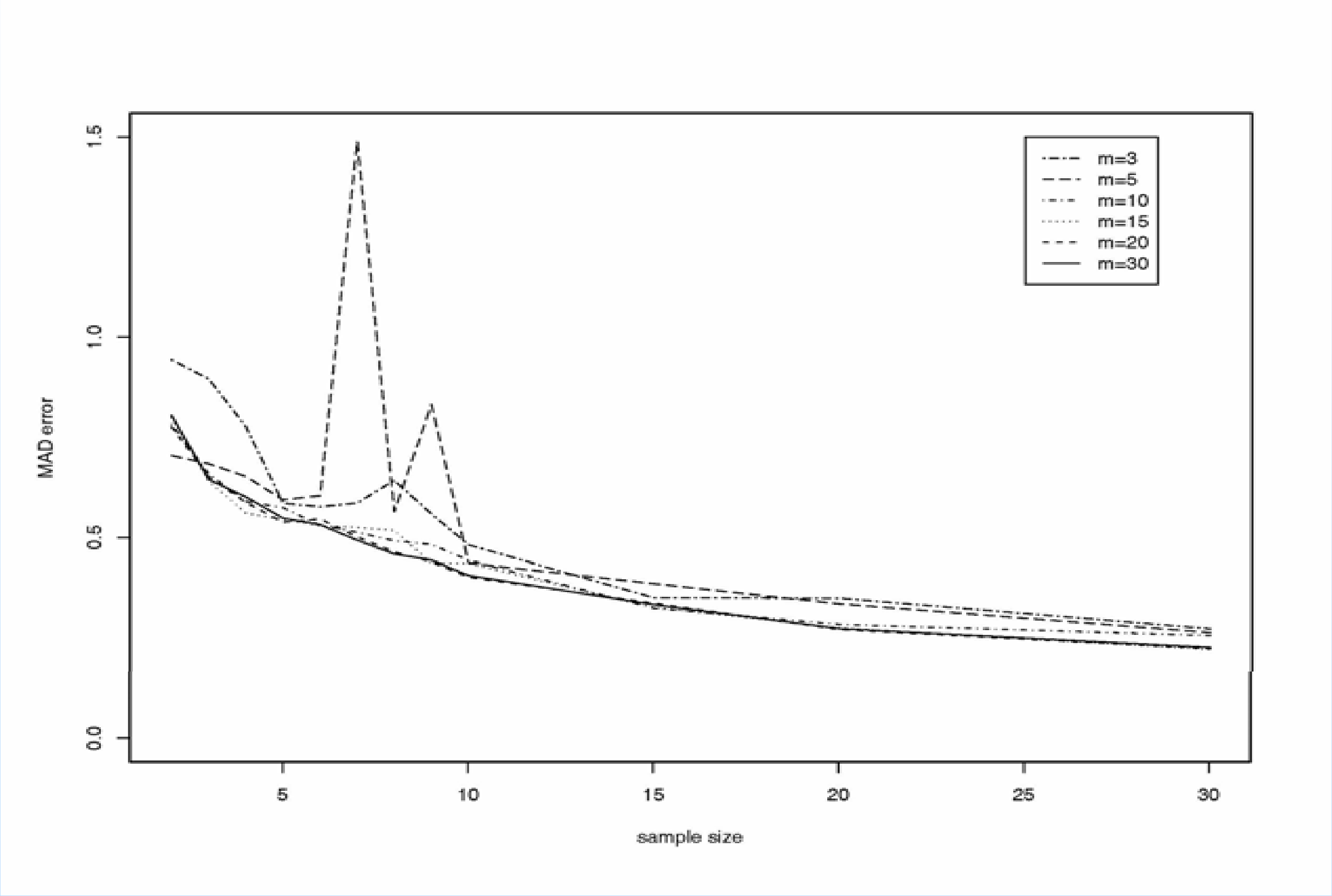
MAD versus MCD

- **CVaR deviation can be observed, while true error in percentile can not be observed in real experiments**
- **MCD provides a conservative estimate of MAD**

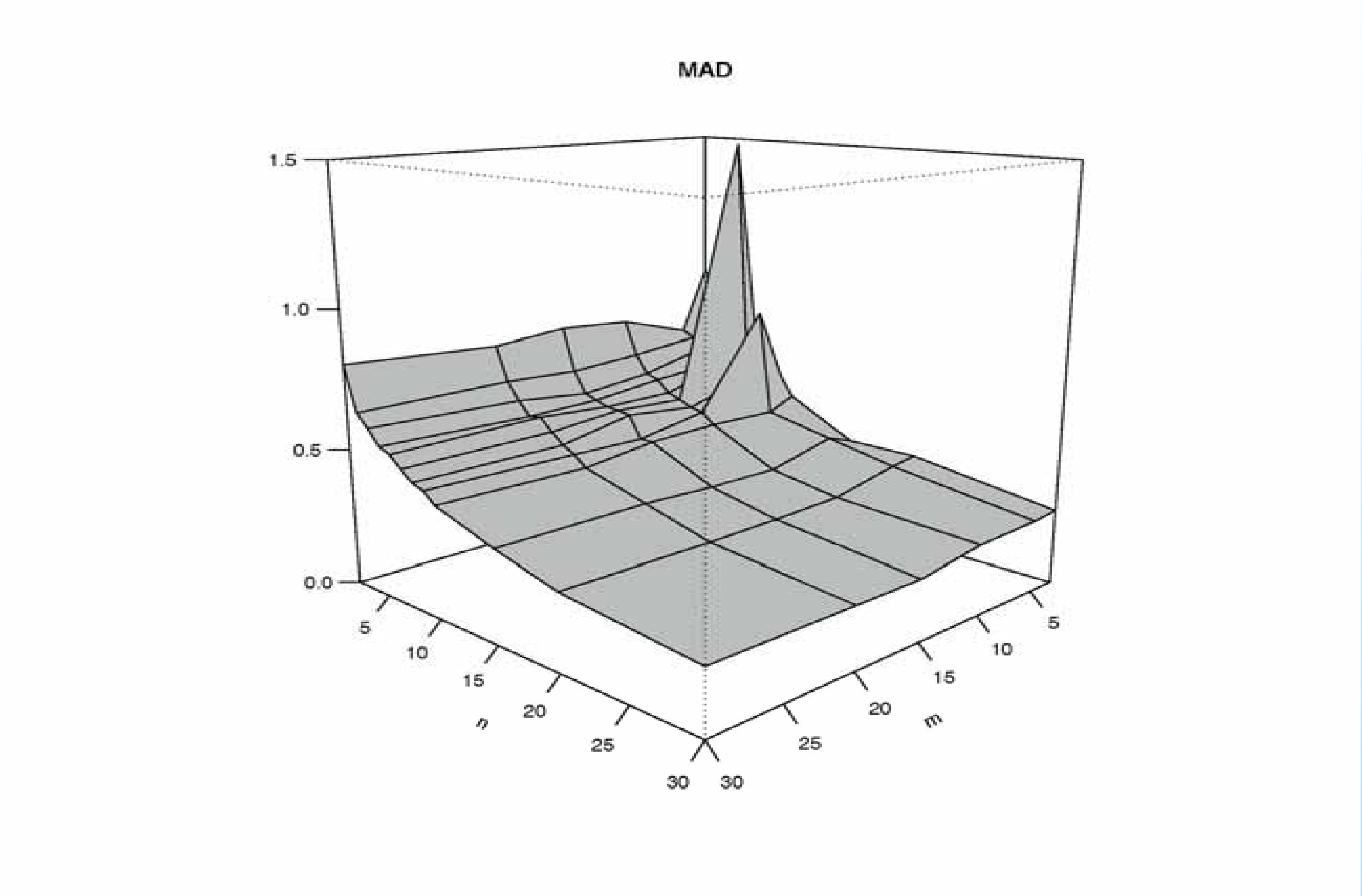


MAD

- **Stable Results for Sample Size n>10,**
- **Unstable Results for m<5 & n<10**

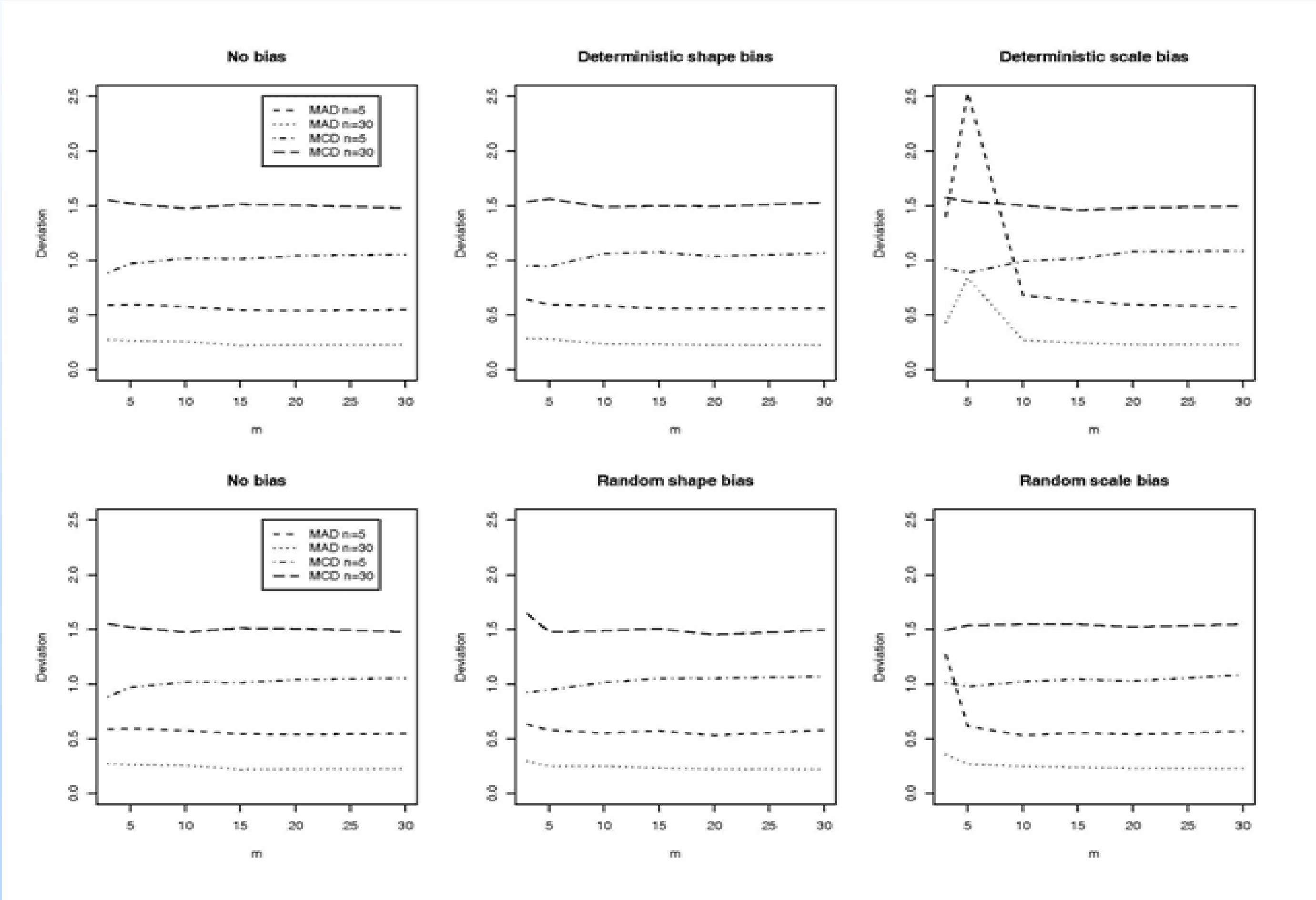


MAD: Stable Results in Wide Range of Parameters, Unstable Results for m<5 & n<10



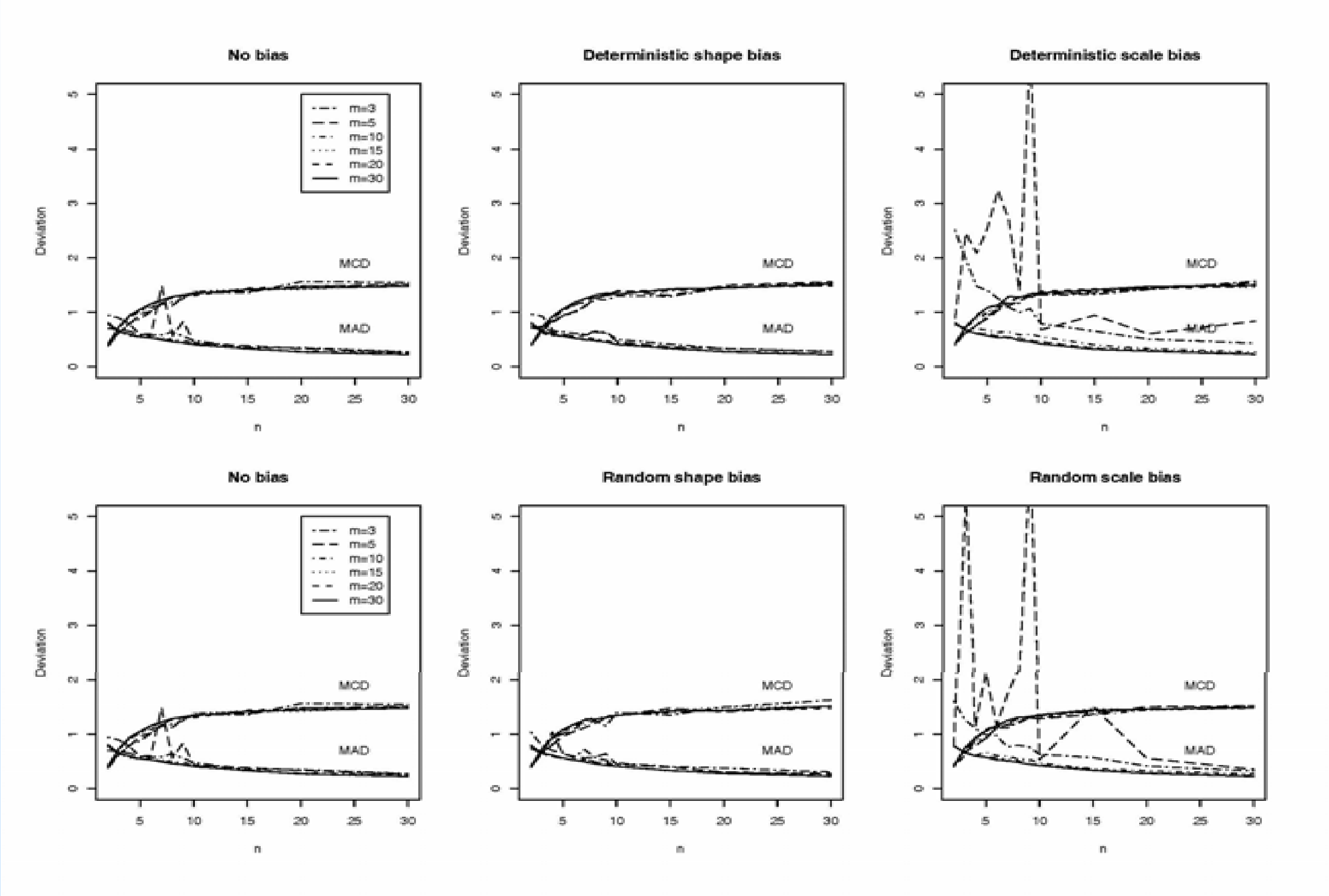
Combining Modeling and Experimental Data: Deterministic Bias and Random Error in Scale and Shape

- **The approach is quite stable to both deterministic and random errors in scale and shape**

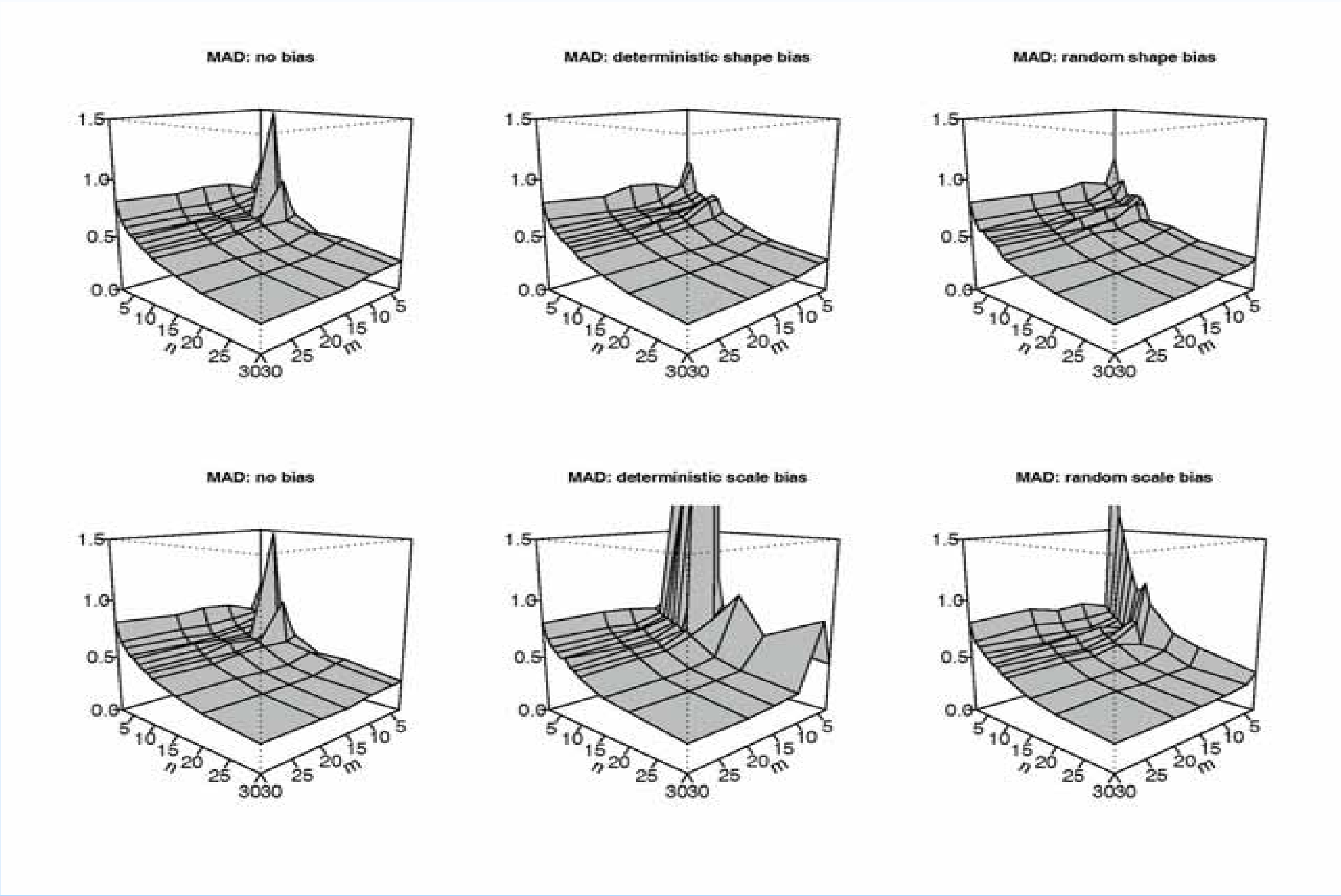


Combining Modeling and Experimental Data: Deterministic Bias and Random Error in Scale and Shape

- **Stable relations between MAD and MCD for various errors in the model (except the case with only 3 stacking sequences)**



Robust Results (MAD): Deterministic and Stochastic Errors in Scale and Shape Parameters



Approved for Public Release, Distribution Unlimited

Recommendations for Implementation

- Certify a model for some area of applications
- Calculate factors for a specific setup of in-sample data and out-of-sample measurements
- Validate the choice of factors with various experiments and Monte Carlo simulations
- Use factor model for estimation of percentiles and allowables in out-of-sample calculations